Microbiology: **Intimate strangers**

Craig Stephens

A more robust view of the diversity of prokaryotes has come from sequencing rRNAs amplified directly from environmental samples. This approach has now been used to examine microbial communities in the human body, revealing populations rich in undescribed species whose impact on humans remains to be determined.

Address: Biology Department, Santa Clara University, Santa Clara, California 95053, USA.
E-mail: cstephens@scu.edu

Current Biology 2000, 10:R272–R275

0960-9822/00/$ – see front matter

An excellent educational series entitled *'Intimate Strangers: Unseen Life on Earth'* was recently aired on public television in the US (see [1]). The series was designed to introduce viewers to the world of microbiology, and emphasized the diversity of microbial life, pointing out many ways in which the activities of microbes in the global biosphere, and their impacts on human health, remain poorly understood. Two papers [2,3], published about the same time as the series was being shown, drive home the concept of

was used to examine bacterial diversity in the subgingival crevice in the mouth [2], and in human feces [3]. The results show that the microbial communities we all carry within us contain many unrecognized bacterial species, perhaps comprising the bulk of these populations. So despite the fact that the human commensal flora has been studied for many decades, we clearly still have a lot to learn about it. Fortunately, this work also suggests approaches to characterizing the structure and dynamic behavior of these microbial communities, and ultimately understanding their ecology.

Microbiologists have long recognized that in most environments, the number of cells visible in the microscope is 10–100-fold greater than the organisms that can actually be isolated as colonies on solid media [4]. The 'great plate count anomaly', as this phenomenon has been referred to, is not due to large numbers of dead cells, as vital dyes or assays of metabolic activity typically reveal that most of these non-recoverable cells are alive [4,5]. Rather, the anomaly seems to be due to a failure to create artificial media and/or incubation conditions that are conducive to the growth of most microbes. Unfortunately, as the identification and characterization of bacterial species has for the last century largely relied on studying the physiological and biochemical — that is, phenotypic — properties of pure cultures, understanding this vast array of microorganisms that cannot be isolated and grown for examination was for a long while nearly impossible.

Pioneering work from the laboratories of Norman Pace and others in the 1980s then began to examine microbial diversity in natural populations by amplifying 16S ribosomal RNA genes directly from environmental samples using the polymerase chain reaction (PCR) [6]. The sequences of these genes were compared to those of known taxa to infer phylogenetic relationships. A truly marvelous diversity of previously unknown prokaryotes, both Bacteria and Archaea, has been revealed by this approach in the last decade [7]. The range of environments being surveyed grows daily, from insect and animal guts to a wide range of soils and sediments, to more exotic locations such as boiling hot springs, superheated ocean vents, subterranean mine shafts, and Antarctic ice caps [4–7]. But to find 'novel' microbes, we need look no further than our own bodies, as was shown early on by application of the ribosomal RNA amplification method to identify the infectious agent in bacillary angiomatosis, a rare disease caused by a bacterial pathogen that had resisted isolation [8].

independent methodology to survey bacteria residing in the subgingival crevice of a healthy person, starting with plaque scraped from the tooth surface below the gum line. Ironically, bacteria from this region were among the first ever observed directly by humans. Over three hundred years ago, Antony von Leeuwenhoek [9] scraped material from between his own teeth, and those of other tolerant townspeople, to examine with his homemade microscopes. To his surprise and pleasure, he saw 'very many small Animals, which moved themselves very extravagantly'. His drawings portray the spirochaetes, cocci and rod-shaped bacteria we now know to be associated with plaque. This microbial community contains many harmless commensal bacteria, but it also has important effects on oral health. Fermentative processes in plaque lead to dental caries, and opportunistic pathogens residing here can cause gingivitis and periodontitis, and contribute to other serious conditions [10]. Kroes *et al.* [2] cultured 28 distinct types of bacteria from their plaque sample, but could assign species designations to fewer than half of them using standard phenotypic methods, reflecting limitations in the current knowledge about many members of this community.

Meanwhile, Suau *et al.* [3] examined bacteria from the opposite end of the gastrointestinal tract. The microbial flora of the intestinal tract has always attracted considerable

Table 1

Comparison of two recent surveys of subgingival plaque and fecal bacterial communities.

	Kroes *et al.* [2]	Suau *et al.* [3]
Starting material	Plaque scraped from the sub-gingival crevice of two teeth of a 39 year old Caucasian male.	Fecal sample from a 40 year old male.
Number of cultivatable species present in sample	28, based on rDNA analysis of 56 phenotypically distinct isolates.	Not reported. Other work suggests < 30 cultivated fecal species per individual based on phenotypic analysis.
% of total bacteria in sample that grew on laboratory media	No estimate provided	21–32%
Basis of rDNA survey	264 clones derived from several libraries were sequenced. 489–690 sequenced bases per clone were compared and used for phylogenetic analysis.	284 clones derived from a single library were sequenced. 400–450 sequenced bases per clone were compared and used for phylogenetic analysis.
Results of rDNA sequence analysis	73 total phylotypes (sequences that differed by <1%) observed; 45 found only by direct amplification, 14 found only by cultivation, and 14 found by both approaches. 36 phylotypes > 99% identical to sequence(s) in ARB database. 37 novel phylotypes < 99% identical to any sequence in ARB database.	82 total 'organizational taxonomic units' (OTUs) – sequences that differed by <2% – observed. 20 OTUs > 98% identical to sequence(s) in ARB database. 62 OTUs < 98% identical to sequence(s) in ARB database.

attention because of its important role in human health. The colon is functionally an anaerobic fermentation chamber packed with bacteria, with cell densities approaching 10^{12} per gram in feces. The resident flora provides some useful functions to its human host, including the synthesis of certain vitamins, and inhibition of intestinal colonization by pathogenic intruders. Suau *et al.* [3] found that only about 30% of the microscopically observable cells in the fecal sample went on to form colonies on agar plates under optimal anaerobic conditions. Although they did not identify the cultivated species in their sample, previous work by Moore and Holdeman [11] using standard phenotypic methods found over 100 bacterial species in cultures from 20 people, with up to 30 cultured species per individual.

DNA extracted from the oral and fecal samples was used as a template for amplification of a portion of the 16S rRNA gene by PCR, using primers specific for bacterial 16S rRNA. The amplification products were inserted into plasmid vectors to generate libraries from which individual clones were randomly selected for DNA sequence determination. More than 200 clones were sequenced in each study, with 400–700 bases from each used in subsequent analyses (Table 1). Sequences greater than 98% identical were considered as a single 'phylotype' likely to be derived from a single species [6]. Both groups found a considerably larger number of phylotypes — 59 oral and 82 fecal — than had ever been isolated from a single person by cultivation. Interested readers should consult the papers [2,3] for further discussion of the criteria used to define phylotypes (Table 1), and the rationales behind them. The consensus among microbial taxonomists is that a bacterial species should be defined in genetic terms as organisms sharing a minimum of 70% overall similarity in genomic DNA sequences, usually as measured by DNA:DNA reassociation [12]. This generally corresponds to differences of at most 2–3% in 16S rRNA sequences between isolates of the same species [12]. (Bacteria often contain multiple rRNA operons in their genome, but these copies typically differ in sequence by less than 1% [2].)

When these phylotypes were compared to a database [13] containing more than 5000 16S rRNA sequences from known bacterial species, plus another 2000 sequences isolated directly from various environments, most were not sufficiently similar to any sequence in the database for it to be likely that they were derived from the same species (Table 1). With the gut bacteria, 62 out of 82 phylotypes appeared to represent novel species. For the plaque specimen, 31 out of 59 phylotypes obtained by direct PCR amplification were novel entities. In contrast, 22 out of 28 of the phylotypes obtained from organisms cultured from plaque material corresponded to known species. The advantage of identifying microbes using rRNA sequence-based methods is evident, as fewer than half of the cultured plaque bacteria were identified at the species level by phenotypic methods. The presence of genetically novel species among even the cultivated plaque bacteria shows that further efforts to catalog the diversity of microorganisms that can be grown in the laboratory would also be useful.

So what are all these 'intimate strangers' doing, toiling in obscurity in our bodies? Phylogenetic relationships inferred from the rDNA sequences suggest some similarities between the plaque and gut flora, probably due to the fact that neither the colon or the subgingival crevice has much oxygen available for respiration, though they do have relatively abundant organic material for fermentation. Obligate anaerobes employing various fermentative pathways have long been known to dominate the culturable fecal flora [11], and anaerobic fermenters have established roles in plaque biology as well [10]. The most abundant among the cloned gut bacterial rDNAs, with 41 novel and 10 previously recognized phylotypes, were members of the *Clostridium* group of spore-forming, Gram-positive obligate anaerobes. This group was also well represented in the gingival specimen, which had nine novel and six previously recognized Clostridial phylotypes. Another major contributor to both samples was the *Bacteroides* group — with 20 fecal and eight gingival phylotypes — consisting of Gram-negative, non-spore forming, obligate anaerobes. In mammalian guts, *Clostridium* and *Bacteroides* species and related genera are undoubtedly major contributors to the anaerobic digestion and fermentation of cellulose and other materials that cannot be broken down by endogenous enzymes in the gastrointestinal tract.

Both studies also found phylotypes from other groups of bacteria in varying numbers. Gingival plaque, for example, had considerably more Streptococci and Actinomycetes than the gut, perhaps because these organisms are relatively tolerant of oxygen. Overall, representatives of four distinct 'divisions' of bacteria were found among the clones sequenced from the fecal sample, and five divisions were found among the plaque-derived clones. 'Division' level groupings have been proposed as the penultimate taxonomic level for bacteria, one level short of the Bacteria–Archaea–Eucarya split [14]. To get a feeling for this evolutionary depth, the 16S rRNA sequences of members of each bacterial division typically differ from other divisions by $> 20\%$, more for example than the divergence between yeast and humans.

Although the number of phylotypes seen in these samples is impressive, it clearly underestimates the actual number of bacterial species in these niches. Statistical analysis of clone distribution estimated that the phylotypes actually found in the fecal library comprise 85% of all the clones in the library, while those seen in the plaque library represent 88% of the clones in the library. Further sequencing might therefore have identified additional phylotypes. Some species frequently observed in these niches were in fact not found among the sequenced clones. Clones representing *Escherichia coli*, the most famous of intestinal bacteria, were not found, but *E. coli* and related facultative anaerobes constitute only about 0.1% of the total bacterial population of the colon, so amplified rDNA from it might not show up among less than 300 sequenced clones.

Bifidobacterium and *Enterococcus* species, on the other hand, are generally prominent components of the human intestinal microflora, but these also were not seen among the clones analyzed. Similarly, no rDNA clones from spirochaetes, such as *Treponema*, were found in the plaque library, even though these usually are abundant in plaque scrapings. Kroes *et al.* [15] point out, as have others, that the composition of rDNA clone libraries can show considerable variation, depending on the PCR primers used. If such surveys are to be reasonably comprehensive and approximate the real distribution of species in a population, careful attention must obviously be paid to sample preparation and the amplification and cloning processes to control for possible over-representation or under-representation of rDNAs from various members of a microbial community.

Of course, these surveys of bacterial diversity beg for follow-up on many fronts. An understanding of the biology of so many new uncultured human commensals at this point will rely largely on inferences from the nearest cultivated phylogenetic neighbors, but for phylotypes only distantly related to known species such comparisons may be of little use. Nevertheless, culture-independent methods can further extend our knowledge of human commensal flora. In an era of high-throughput DNA sequencing, analyzing more clones from libraries generated using carefully chosen primers that minimize amplification biases should be a straightforward way of finding more species. Examining libraries generated from several people is important to expand the catalog of species found in these niches, and begin to address the question of how the microbial flora might vary between people [16].

Other places in the human body with important commensal flora should be explored too — the nasopharyngeal region and the urogenital tract come to mind, as each has a distinct, rich flora that includes opportunistic pathogens. Once the range of organisms normally present in each niche has been ascertained, quantitative methods for assessing population composition can be applied using oligonucleotide probes designed from rRNA sequences. For example, it might be possible for the composition of a sample to be rapidly evaluated by microarray analysis using species-specific or group-specific probes, without ever culturing bacteria in the laboratory. Assessments of microbial population dynamics during the course of an infection, as a consequence of antibiotic treatment, or as a function of age, diet, ethnicity, or any number of other variables would almost certainly have significant relevance to improving human health.

Acknowledgements
The work of the author is supported by National Science Foundation grant MCB-9808798.

References

1. Annenberg/CPB Projects web site: http://www.learner.org/collections/multimedia/science/isseries
2. Kroes I, Lepp P, Relman D: **Bacterial diversity within the human subgingival crevice.** *Proc Natl Acad Sci USA* 1999, **96:**14547-14552.
3. Suau A, Bonnet R, Sutren M, Godon J, Gibson G, Collins M, Dore J: **Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut.** *Appl Environ Micro* 1999, **65:**4799-4807.
4. Amman R, Ludwig W, Schleifer K-H: **Phylogenetic identification and in situ detection of individual microbial cells without cultivation.** *Microbiol Rev* 1995, **59:**143-169.
5. Staley J, Konopka A: **Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats.** *Annu Rev Microbiol* 1985, **39:**321-346.
6. Pace N: **A molecular view of microbial diversity and the biosphere.** *Science* 1997, **276:**734-740.
7. Woese C: **Whither microbiology? Phylogenetic trees.** *Curr Biol* 1996, **6:**1060-1063.
8. Relman D, Loutit J, Schmidt T, Falkow S, Tompkins L: **The agent of bacillary angiomatosis. An approach to the identification of uncultured pathogens.** *N Engl J Med* 1990, **323:**1573-1580
9. von Leewenhoeck A: **Microscopical observations about animals in the scurf of the teeth.** *Phil Trans Royal Soc Lon* 1684, **14:**568-574.
10. Meyer D, Fives-Taylor P: **Oral pathogens: from dental plaque to cardiac disease.** *Curr Opin Microbiol* 1998, **1:**88-95.
11. Moore W, Holdeman L: **Human fecal flora: the normal flora of 20 Japanese-Hawaiians.** *Appl Microbiol* 1974, **27:**961-979.
12. Stackebrandt E, Goebel B: **Taxonomic note: a place for DNA reassociation and 16S rRNA in the present species definition in bacteriology.** *Intl J Syst Bacteriol* 1994, **44:**846-849.
13. Maidak B, Cole J, Parker C, Garrity G, Larsen N, Li B, Lilburn T, McCaughey M, Olsen G, Overbeek R, *et al.*: **A new version of the RDP (Ribosomal Database Project).** *Nucleic Acids Res* 1999, **27:**171-173.
14. Hugenholtz P, Goebel B, Pace N: **Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity.** *J Bacteriol* 1998, **180:**4765-4774.
15. Wilson KH, Blitchington RB: **Human colonic biota studied by ribosomal DNA sequence analysis.** *Appl Environ Microbiol* 1996, **62:**2273-2278.
16. Zoetendal E, Akkermans A, De Vos W: **Temperature gradient gel electrophoresis analysis of 16S rRNA from human fecal samples reveals stable and host-specific communities of active bacteria.** *Appl Environ Microbiol* 1998, **64:**3854-3859.